

R. Swartz

1641

#6
MB
06/09/00

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/461,774

DATE: 05/19/2000
TIME: 18:58:25

Input Set: I461774.RAW

This Raw Listing contains the General Information
Section and up to first 5 pages.

ENTERED

RECEIVED
JUN 08 2000

1 <110> APPLICANT: CHAN, Lily
2 CHUNG, Maxey Ching Ming
3 LIM, Renee Lay Hong
4 <120> TITLE OF INVENTION: Bacterial-derived molecules and therapeutic and
5 diagnostic uses therefor
6 <130> FILE REFERENCE: 1781-0180P
7 <140> CURRENT APPLICATION NUMBER: US/09/461,774
8 <141> CURRENT FILING DATE: 1999-12-15
9 <160> NUMBER OF SEQ ID NOS: 22
10 <170> SOFTWARE: PatentIn Ver. 2.1
11 <210> SEQ ID NO 1
12 <211> LENGTH: 1617
13 <212> TYPE: DNA
14 <213> ORGANISM: Mycobacterium tuberculosis
15 <220> FEATURE:
16 <221> NAME/KEY: CDS
17 <222> LOCATION: (1)..(1617)
18 <220> FEATURE:
19 <221> NAME/KEY: mat_peptide
20 <222> LOCATION: (1)..(1617)
21 <400> SEQUENCE: 1
22 agc aag ctg atc gaa tac gac gaa acc gcg cgt cgc gcc atg gag gtc 48
23 Ser Lys Leu Ile Glu Tyr Asp Glu Thr Ala Arg Arg Ala Met Glu Val
24 1 5 10 15
25 ggc atg gac aag ctg gcc gac acc gtg cgg gtg acg ctg ggg ccg cgc 96
26 Gly Met Asp Lys Leu Ala Asp Thr Val Arg Val Thr Leu Gly Pro Arg
27 20 25 30
28 ggc cgg cat gtg gtg ctg gcc aag gcg ttt ggc gga ccc acg gtt acc 144
29 Gly Arg His Val Val Leu Ala Lys Ala Phe Gly Gly Pro Thr Val Thr
30 35 40 45
31 aac gac ggc gtc acg gtg gca cgt gag atc gag ctg gaa gat ccg ttt 192
32 Asn Asp Gly Val Thr Val Ala Arg Glu Ile Glu Leu Glu Asp Pro Phe
33 50 55 60
34 gaa gac ttg ggc gcc cag ctg gtg aag tcg gtg gcc acc aag acc aac 240
35 Glu Asp Leu Gly Ala Gln Leu Val Lys Ser Val Ala Thr Lys Thr Asn
36 65 70 75 80
37 gat gtt gcc ggt gac ggc acc acc acc gca acc atc ttg gcg cag gca 288
38 Asp Val Ala Gly Asp Gly Thr Thr Thr Ala Thr Ile Leu Ala Gln Ala
39 85 90 95
40 ctg atc aag ggc ggc ctg agg cta gtg gcc gcc ggc gtc aac ccg gtc 336
41 Leu Ile Lys Gly Gly Leu Arg Leu Val Ala Ala Gly Val Asn Pro Val
42 100 105 110
43 gcg ctc ggc gtg gga atc ggc aag gcc gcc gac gcg gta ttc gag gcg 384
44 Ala Leu Gly Val Gly Ile Gly Lys Ala Ala Asp Ala Val Phe Glu Ala

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/461.774

DATE: 05/19/2000
TIME: 18:58:25

Input Set: **I461774.RAW**

45																		
46	ctg	ctg	gca	tcg	gcc	acg	ccg	gtg	tcc	ggc	aag	acc	ggc	atc	gcg	cag	432	
47	Leu	Leu	Ala	Ser	Ala	Thr	Pro	Val	Ser	Gly	Lys	Thr	Gly	Ile	Ala	Gln		
48		130						135					140					
49	gtg	gcg	acg	gtg	tcc	tcg	cgc	gac	gag	cag	atc	ggt	gac	ctg	gtt	ggc	480	
50	Val	Ala	Thr	Val	Ser	Ser	Arg	Asp	Glu	Gln	Ile	Gly	Asp	Leu	Val	Gly		
51	145						150					155				160		
52	gaa	gcg	atg	aac	aag	gtc	ggc	cac	gac	agc	gtg	gtc	agc	gtc	aag	gaa	528	
53	Glu	Ala	Met	Asn	Lys	Val	Gly	His	Asp	Ser	Val	Val	Ser	Val	Lys	Glu		
54					165					170					175			
55	tcc	tcc	acg	ctg	ggc	acc	gag	ttg	gag	ttc	acc	gaa	ggt	att	ggc	ttc	576	
56	Ser	Ser	Thr	Leu	Gly	Thr	Glu	Leu	Glu	Phe	Thr	Glu	Gly	Ile	Gly	Phe		
57			180						185					190				
58	cac	aaa	ggc	ttc	ttg	tcg	gca	tac	ttc	gtt	acc	gac	ttc	gat	aac	cag	624	
59	His	Lys	Gly	Phe	Leu	Ser	Ala	Tyr	Phe	Val	Thr	Asp	Phe	Asp	Asn	Gln		
60		195						200					205					
61	cag	gcg	gtg	ctc	gag	gac	gcg	ttg	atc	ctg	cta	cac	caa	gac	aag	atc	672	
62	Gln	Ala	Val	Leu	Glu	Asp	Ala	Leu	Ile	Leu	Leu	His	Gln	Asp	Lys	Ile		
63		210					215					220						
64	agc	tcg	ctt	ccc	gat	ctg	ttg	cca	ttg	ctg	gaa	aag	gtt	gca	gga	acg	720	
65	Ser	Ser	Leu	Pro	Asp	Leu	Leu	Pro	Leu	Leu	Glu	Lys	Val	Ala	Gly	Thr		
66	225					230					235					240		
67	ggt	aag	cca	cta	ctg	atc	gtg	gct	gaa	gac	gtg	gag	ggc	gaa	gcg	ttg	768	
68	Gly	Lys	Pro	Leu	Leu	Ile	Val	Ala	Glu	Asp	Val	Glu	Gly	Glu	Ala	Leu		
69				245						250					255			
70	gcg	acg	ctg	gtc	gtc	aac	gcg	att	cgc	aag	acg	ttg	aaa	gcg	gtc	gcg	816	
71	Ala	Thr	Leu	Val	Val	Asn	Ala	Ile	Arg	Lys	Thr	Leu	Lys	Ala	Val	Ala		
72			260						265					270				
73	gtc	aag	ggg	ccg	tac	ttc	ggt	gac	cgc	cgt	aag	gcg	ttc	ctt	gag	gac	864	
74	Val	Lys	Gly	Pro	Tyr	Phe	Gly	Asp	Arg	Arg	Lys	Ala	Phe	Leu	Glu	Asp		
75		275						280				285						
76	ctg	gcg	gtg	gtg	acg	ggt	ggc	cac	gtg	gtc	aac	ccc	gac	gcc	ggc	att	912	
77	Leu	Ala	Val	Val	Thr	Gly	Gly	His	Val	Val	Asn	Pro	Asp	Ala	Gly	Ile		
78		290					295					300						
79	gtg	ctg	cgc	gag	gtg	ggc	ttg	gag	gtg	ctg	ggc	tcg	gcc	cga	cgc	gtg	960	
80	Val	Leu	Arg	Glu	Val	Gly	Leu	Glu	Val	Leu	Gly	Ser	Ala	Arg	Arg	Val		
81	305					310												

PAGE: 3

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/461,774

 DATE: 05/19/2000
 TIME: 18:58:25

Input Set: I461774.RAW

```

95   Ala Leu Lys Glu Arg Lys Glu Ser Val Glu Asp Ala Val Ala Ala Ala
96   385                               390                               395                               400
97   aag gcc gcg gtc gag gag ggc atc gtc cct ggt ggg gga gcc tcg ctc   1248
98   Lys Ala Ala Val Glu Glu Gly Ile Val Pro Gly Gly Gly Ala Ser Leu
99                               405                               410                               415
100  atc cac cag gcc cgc aag gcg ctg acc gaa ctg cgt gcg tcg ctg acc   1296
101  Ile His Gln Ala Arg Lys Ala Leu Thr Glu Leu Arg Ala Ser Leu Thr
102                               420                               425                               430
103  ggt gac gag gtc ctc ggt gtc gac gtg ttc tcc gaa gcc ctt gcc gcg   1344
104  Gly Asp Glu Val Leu Gly Val Asp Val Phe Ser Glu Ala Leu Ala Ala
105                               435                               440                               445
106  ccg ttg ttc tgg atc gcc gcc aac gct ggc ttg gac ggc tcg gtg gtg   1392
107  Pro Leu Phe Trp Ile Ala Ala Asn Ala Gly Leu Asp Gly Ser Val Val
108                               450                               455                               460
109  gtc aac aag gtc agc gag cta ccc gcc ggg cat ggg ctg aac gtg aac   1440
110  Val Asn Lys Val Ser Glu Leu Pro Ala Gly His Gly Leu Asn Val Asn
111  465                               470                               475                               480
112  acc ctg agc tat ggt gac ttg gcc gct gac ggc gtc atc gac ccg gtc   1488
113  Thr Leu Ser Tyr Gly Asp Leu Ala Ala Asp Gly Val Ile Asp Pro Val
114                               485                               490                               495
115  aag gtg act agg tcg gcg gtg ttg aac gcg tca tcg gtt gcc cgg atg   1536
116  Lys Val Thr Arg Ser Ala Val Leu Asn Ala Ser Ser Val Ala Arg Met
117                               500                               505                               510
118  gta ctc acc acc gag acg gtc gtg gtc gac aag ccg gcc aag gca gaa   1584
119  Val Leu Thr Thr Glu Thr Val Val Val Asp Lys Pro Ala Lys Ala Glu
120                               515                               520                               525
121  gat cac gac cat cac cac ggg cac gcg cac tga
122  Asp His Asp His His His Gly His Ala His
123  530                               535
124  <210> SEQ ID NO 2
125  <211> LENGTH: 538
126  <212> TYPE: PRT
127  <213> ORGANISM: Mycobacterium tuberculosis
128  <400> SEQUENCE: 2
129  Ser Lys Leu Ile Glu Tyr Asp Glu Thr Ala Arg Arg Ala Met Glu Val
130  1                               5                               10                               15
131  Gly Met Asp Lys Leu Ala Asp Thr Val Arg Val Thr Leu Gly Pro Arg
132  20                               25                               30
133  Gly Arg His Val Val Leu Ala Lys Ala Phe Gly Gly Pro Thr Val Thr
134  35                               40                               45
135  Asn Asp Gly Val Thr Val Ala Arg Glu Ile Glu Leu Glu Asp Pro Phe
136  50                               55                               60
137  Glu Asp Leu Gly Ala Gln Leu Val Lys Ser Val Ala Thr Lys Thr Asn
138  65                               70                               75                               80
139  Asp Val Ala Gly Asp Gly Thr Thr Thr Ala Thr Ile Leu Ala Gln Ala
140  85                               90                               95
141  Leu Ile Lys Gly Gly Leu Arg Leu Val Ala Ala Gly Val Asn Pro Val
142  100                              105                              110
143  Ala Leu Gly Val Gly Ile Gly Lys Ala Ala Asp Ala Val Phe Glu Ala
144  115                              120                              125

```

RECEIVED

JUN 02 2000

TECH CENTER 1600/2300

PAGE: 4

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/461,774

 DATE: 05/19/2000
 TIME: 18:58:25

Input Set: I461774.RAW

145	Leu	Leu	Ala	Ser	Ala	Thr	Pro	Val	Ser	Gly	Lys	Thr	Gly	Ile	Ala	Gln
146		130					135					140				
147	Val	Ala	Thr	Val	Ser	Ser	Arg	Asp	Glu	Gln	Ile	Gly	Asp	Leu	Val	Gly
148	145					150					155					160
149	Glu	Ala	Met	Asn	Lys	Val	Gly	His	Asp	Ser	Val	Val	Ser	Val	Lys	Glu
150					165					170					175	
151	Ser	Ser	Thr	Leu	Gly	Thr	Glu	Leu	Glu	Phe	Thr	Glu	Gly	Ile	Gly	Phe
152				180						185				190		
153	His	Lys	Gly	Phe	Leu	Ser	Ala	Tyr	Phe	Val	Thr	Asp	Phe	Asp	Asn	Gln
154			195					200					205			
155	Gln	Ala	Val	Leu	Glu	Asp	Ala	Leu	Ile	Leu	Leu	His	Gln	Asp	Lys	Ile
156		210					215					220				
157	Ser	Ser	Leu	Pro	Asp	Leu	Leu	Pro	Leu	Leu	Glu	Lys	Val	Ala	Gly	Thr
158	225					230					235					240
159	Gly	Lys	Pro	Leu	Leu	Ile	Val	Ala	Glu	Asp	Val	Glu	Gly	Glu	Ala	Leu
160				245						250					255	
161	Ala	Thr	Leu	Val	Val	Asn	Ala	Ile	Arg	Lys	Thr	Leu	Lys	Ala	Val	Ala
162				260						265				270		
163	Val	Lys	Gly	Pro	Tyr	Phe	Gly	Asp	Arg	Arg	Lys	Ala	Phe	Leu	Glu	Asp
164			275					280					285			
165	Leu	Ala	Val	Val	Thr	Gly	Gly	His	Val	Val	Asn	Pro	Asp	Ala	Gly	Ile
166		290					295				300					
167	Val	Leu	Arg	Glu	Val	Gly	Leu	Glu	Val	Leu	Gly	Ser	Ala	Arg	Arg	Val
168	305					310					315					320
169	Val	Val	Ser	Lys	Asp	Asp	Thr	Val	Ile	Val	Asp	Gly	Gly	Gly	Thr	Ala
170				325						330					335	
171	Glu	Ala	Val	Ala	Asn	Arg	Ala	Asn	His	Leu	Arg	Ala	Glu	Ile	Asp	Lys
172				340					345					350		
173	Ser	Asp	Ser	Asp	Trp	Asp	Arg	Glu	Lys	Leu	Gly	Glu	Arg	Leu	Ala	Lys
174			355					360					365			
175	Leu	Ala	Gly	Gly	Val	Ala	Val	Ile	Lys	Val	Gly	Ala	Ala	Thr	Asp	Thr
176		370					375					380				
177	Ala	Leu	Lys	Glu	Arg	Lys	Glu	Ser	Val	Glu	Asp	Ala	Val	Ala	Ala	Ala
178	385					390					395					400
179	Lys	Ala	Ala	Val	Glu	Glu	Gly	Ile	Val	Pro	Gly	Gly	Gly	Ala	Ser	Leu
180				405						410					415	
181	Ile	His	Gln	Ala	Arg	Lys	Ala	Leu	Thr	Glu	Leu	Arg	Ala	Ser	Leu	Thr
182				420						425				430		
183	Gly	Asp	Glu	Val	Leu	Gly	Val	Asp	Val	Phe	Ser	Glu	Ala	Leu	Ala	Ala
184			435					440					445			
185	Pro	Leu	Phe	Trp	Ile	Ala	Ala	Asn	Ala	Gly	Leu	Asp	Gly	Ser	Val	Val
186		450					455					460				
187	Val	Asn	Lys	Val	Ser	Glu	Leu	Pro	Ala	Gly	His	Gly	Leu	Asn	Val	Asn
188	465					470					475					480
189	Thr	Leu	Ser	Tyr	Gly	Asp	Leu	Ala	Ala	Asp	Gly	Val	Ile	Asp	Pro	Val
190				485						490					495	
191	Lys	Val	Thr	Arg	Ser	Ala	Val	Leu	Asn	Ala	Ser	Ser	Val	Ala	Arg	Met
192				500						505				510		
193	Val	Leu	Thr	Thr	Glu	Thr	Val	Val	Val	Asp	Lys	Pro	Ala	Lys	Ala	Glu
194				515					520					525		

PAGE: 5

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/461,774

 DATE: 05/19/2000
 TIME: 18:58:25

Input Set: I461774.RAW

```

195      Asp His Asp His His His Gly His Ala His
196          530                      535
197 <210> SEQ ID NO 3
198 <211> LENGTH: 1560
199 <212> TYPE: DNA
200 <213> ORGANISM: Mycobacterium tuberculosis
201 <220> FEATURE:
202 <221> NAME/KEY: CDS
203 <222> LOCATION: (1)..(1560)
204 <220> FEATURE:
205 <221> NAME/KEY: mat_peptide
206 <222> LOCATION: (1)..(1560)
207 <400> SEQUENCE: 3
208      gcg gcc atg tgg cgc cgc aga ccg ttg agc tcg gcg ctg ctg tcc ttc      48
209      Ala Ala Met Trp Arg Arg Arg Pro Leu Ser Ser Ala Leu Leu Ser Phe
210          1                      5                      10                      15
211      ggg ttg ctg ctc ggc gga ctg ctc cta gca gcg ccc ccg ttg gcc ggc      96
212      Gly Leu Leu Leu Gly Gly Leu Leu Leu Ala Ala Pro Pro Leu Ala Gly
213          20                      25                      30
214      gcg act gaa gaa ccc ggc gcc ggc caa acc ccg ggt gcg ccg gtc gtg      144
215      Ala Thr Glu Glu Pro Gly Ala Gly Gln Thr Pro Gly Ala Pro Val Val
216          35                      40                      45
217      gcg ccg caa cag agt tgg aac agc tgc cgc gag ttc atc gcc gac acc      192
218      Ala Pro Gln Gln Ser Trp Asn Ser Cys Arg Glu Phe Ile Ala Asp Thr
219          50                      55                      60
220      agc gaa att cgc act gca cgc tgc gcg acg gtg tcc gtc ccc gtc gac      240
221      Ser Glu Ile Arg Thr Ala Arg Cys Ala Thr Val Ser Val Pro Val Asp
222          65                      70                      75                      80
223      tac gac caa ccc ggt ggg aca caa gcg aag ttg gcg gtg atc cgc gtc      288
224      Tyr Asp Gln Pro Gly Gly Thr Gln Ala Lys Leu Ala Val Ile Arg Val
225          85                      90                      95
226      ccc gcg acg gga cag cga ttc gga gca ctg ctg gtc aat cct ggg gga      336
227      Pro Ala Thr Gly Gln Arg Phe Gly Ala Leu Leu Val Asn Pro Gly Gly
228          100                     105                     110
229      ccc ggg gcg tcg gcg gtc gac atg gtc gcc gct atg gca ccc gcg atc      384
230      Pro Gly Ala Ser Ala Val Asp Met Val Ala Ala Met Ala Pro Ala Ile
231          115                     120                     125
232      gcc gac acc gac att ctc cgc cac ttc gac ctg gtg ggc ttc gac ccg      432
233      Ala Asp Thr Asp Ile Leu Arg His Phe Asp Leu Val Gly Phe Asp Pro
234          130                     135                     140
235      aga ggg gtc ggc cac tcg acc cct gcg ttg cgg tgt cgc acc gac gcc      480
236      Arg Gly Val Gly His Ser Thr Pro Ala Leu Arg Cys Arg Thr Asp Ala
237          145                     150                     155                     160
238      gag ttc gac gcg tac cgg cgc gat ccg atg gcc gac tac agt ccg gcc      528
239      Glu Phe Asp Ala Tyr Arg Arg Asp Pro Met Ala Asp Tyr Ser Pro Ala
240          165                     170                     175
241      ggc gtc acc cac gtc gaa cag gtc tac cgg cag ttg gcc cag gac tgt      576
242      Gly Val Thr His Val Glu Gln Val Tyr Arg Gln Leu Ala Gln Asp Cys
243          180                     185                     190
244      gtt gac cgg atg ggc ttc agc ttc ttg gcc aat atc ggt acc gcg tcc      624

```

PAGE: 6

VERIFICATION SUMMARY
PATENT APPLICATION US/09/461,774

DATE: 05/19/2000

TIME: 18:58:25

Input Set: I461774.RAW

Line ? Error/Warning

Original Text

306 W Invalid/Missing Amino Acid Numbering

515

520